SciDAC Visualization and Analytics Center for Enabling Technologies
Project Management Plan
Version 1.0

E. Wes Bethel [*]and Chris Johnson [†]
Principal Investigators

Charles Hansen, Claudio Silva, Steven Parker, Allen Sanderson,
Xavier Tricoche, Martin Cole [‡]

Sean Ahern, George Ostrouchov, Jeremy Meredith [§]

Valerio Pascucci, Hank Childs, Peer-Timo Bremer, Kathleen Bonnell [¶]

Ken Joy, Bernd Hamann [‖]

November 2006

[*]Lawrence Berkeley National Laboratory
[†]Scientific Computing Institute, University of Utah
[‡]Scientific Computing Institute, University of Utah
[§]Oak Ridge National Laboratory
[¶]Lawrence Livermore National Laboratory
[‖]University of California – Davis

Contents

1 Introduction

This SciDAC2 Center for Enabling Technology focuses on leveraging scientific visualization and analytics software technology as an enabling technology for increasing scientific productivity and insight. Advances in computational technology have resulted in an "information big bang," which in turn has created a significant data understanding challenge. This challenge is widely acknowledged to be one of the primary bottlenecks in contemporary science. The vision for our Center is to respond directly to that challenge by adapting, extending, creating when necessary and deploying visualization and data understanding technologies for our science stakeholders. Using an organizational model as a Visualization and Analytics Center for Enabling Technologies (VACET), we are well positioned to be responsive to the needs of a diverse set of scientific stakeholders in a coordinated fashion using a range of visualization, mathematics, statistics, computer and computational science and data management technologies.

Our team consists of international leaders in scientific visualization and analysis with a strong record of creating and deploying visualization software and of effectively collaborating with application stakeholders. Our visualization team already has strong relationships with application scientists, and team members holding positions within DOE will ensure our solutions are deployed at DOE's large-scale open computing facilities, including the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory and the Leadership Computing Facility (LCF) at Oak Ridge National Laboratory. Most members of our team have been active in conducting visualization research, development, and deployment activities during the first five years of SciDAC. We are collectively dedicated to our Center's mission: to foster scientific insight through creating and deploying effective data understanding technology that is truly responsive to the needs of the scientific research community.

VACET officially "began business" at the start of 1 October 2006, or the beginning of FY07. As of the time of this writing – a little over a month later – we are very much in the "startup phase" of operations. While this document is the first in a series of "Project Management Plans" that document ongoing VACET operations, this first version is best viewed as a combination of a preliminary status report and initial reporting of time-appropriate elements of a more formal "Project Management Plan."

Our project's current "startup operations" focus on requirements gathering and analysis, which will then provide the material needed for deriving a more formal work plan containing milestones and deliverables. As this document details, we have been very active in engaging stakeholders over the first month of operations. This activity, which is the first step in our project, aims to identify stakeholder needs from about a dozen different projects. The next step is to finalize and vet the list of needs with stakeholders, translate the needs into work tasks with estimates of time required to complete the tasks, and assign work tasks to team members. Therefore, while this Project Management Plan is best viewed as being very preliminary in breadth and depth, it accurately conveys a high degree of productivity at this early stage of the project.

The remaining sections of this document are organized as follows. Section 2 discusses "baseline activities," which indicate where we are now, where we want to be and how we'll get there. Section 3 outlines the roles of key project personnel. Section 4 discusses our software development, maintenance and ongoing support plan along with intellectual property issues and their resolution. Section 5 outlines our strategy for communication and interactions within VACET, other SciDAC Science Applications, Centers and Institutes, as well as outreach. Section 6 presents a brief statement about the differences between the Visualization Center and Institute. Section 7 itemizes known risks, their severity, impact to the project and remediation strategies. Section 8 contains additional project management materials.

2 Baseline Activities

2.1 Current Status – Fall 2006

VACET officially "began business" at the start of 1 October 2006, or the beginning of FY07. As of the time of this writing – about a month later – we are very much in the "startup phase" of operations.

At the time we created our VACET proposal – late in 2005 and early 2006 – we engaged in discussion with several potential science application stakeholders for the purpose of documenting their data understanding needs, translating those needs into actionable items, prioritizing them and then presenting them in our proposal. Our proposal's scope, while very much reflecting state-of-the-art in scientific visualization and analytics, was very much driven by the needs of our potential stakeholders.

One difficulty with being a "startup project" is that we were not certain exactly whom our customers would be until late in FY06. All of our stakeholders are themselves SciDAC projects, many of whom are new, "startup projects." The net effect of this situation is that we and our stakeholders all officially "began business" at the start of FY06. Despite this timing, VACET is presently in a very good position due to the strong personal and professional relationships we have with our science stakeholder projects.

In the sections that follow, we outline project objectives that span three different time horizons. The three different horizons reflect VACET's growth from "a startup" to a healthy, vibrant and productive organization. The first subsection identifies activities to-date and near-term objectives. The following two sketch high-level objectives farther out in time.

2.2 Near-Term Objectives – Fall 2007

2.2.1 Identify Stakeholders and Needs

As most project work tasks and deliverables are aimed at deploying software that meets the needs of science stakeholders, the foremost VACET objectives at the time of this writing are: (1) to identify application science stakeholders; (2) to formally document their needs, including having the stakeholders participate in a deliverable prioritization and vetting process; (3) to derive a formalized work plan, milestones and deliverables that result in those needs/objectives being met.

To date, VACET team members have participated in many such activities, including VACET representation at kickoff meetings for other SciDAC projects (APDEC, Astrophysics), conference calls and face-to-face meetings with stakeholders (Accelerator, Astrophysics, Fusion, Climate, Combustion, Math/APDEC, CS/SDM) and email exchanges with numerous stakeholders. The results of these activities include a preliminary set of stakeholder needs that is currently undergoing a revision and vetting process.

One theme that emerges from these VACET-stakeholder interactions is the desire for VACET to participate in the role as "team member" rather than as a "disconnected collaborator." We envision most, if not all, of our stakeholder relationships will be characterized as VACET acting as a team member. Such a level of interaction will likely require a significant investment of VACET time. We include such time requirements in our planning process.

2.2.2 Accelerator Modeling

We have had several face-to-face and email conversations with Rob Ryne, head of the "Advanced Computing for 21st Century Accelerator Science and Technology" project and with Cameron Geddes, PI on the 2006 INCITE project entitled "Particle-In-Cell Simulation of Laser Wakefield Particle Acceleration." We are using their input to gain a "head start" on documenting and prioritizing needs to be prepared when the Accelerator SciDAC proposal is submitted later in FY06. Generally speaking, they have expressed interest in several related topics: high performance I/O, community-wide data models and analysis tools, the ability to perform simultaneous 2D and 3D visual data analysis of time-varying accelerator simulation data consisting of both particles and fields. In particular, this project needs to be able to quickly identify particles that exhibit specific properties, such as those that form halos or diverge from the main beamline and then trace them forwards and backwards in time as well as study correlations between properties of select particles and field variables. This group has also expressed interest in domain-specific visual analysis capabilities like "frequency maps" computed from many timesteps that show spectral characteristics of a beamline cross-section and how it evolves over time.

2.2.3 Astrophysics

VACET team members attended the kickoff meeting of Stan Woosley's Computational Astrophysics Consortium SciDAC Science Application held at Stanford on 13 October 2006. We gave a brief overview of VACET's objectives and of VACET's visualization technology. We used the opportunity to enumerate our understanding of the needs of the computational and experimental astrophysics community; we gathered these requirements through a series of interviews over the past eight months. Those needs include: (1) visualization tools capable of displaying data from AMR grids that have tremendous spatial and temporal scale; (2) effective visualization techniques for radiation flux; (3) new grid types; (4) feature identification and tracking, where features could be a function of combinations of data ranges or based upon topology; (5) feature mining and visual correlation; (6) effective remote visualization capability; (7) transparent access to distributed data; (8) simultaneous display of particle- and field-based data; (9) the ability to perform quantitative simulation-simulation and simulation-observation comparative visual analysis.

Woosley indicates that J. Bell's effort would comprise the bulk of the new code development, and therefore VACET should focus primarily on working with Bell. We have had detailed conversations with Bell and his group to create preliminary, priority-ordered list of needs that is currently undergoing a process of vetting and transformation into a VACET task list.

2.2.4 Climate

VACET has had interactions with two different SciDAC climate projects: (1) the Earth Systems Grid (D. Williams, LLNL), and (2) Modeling the Earth System (J. Drake, ORNL). These two projects are complementary in the sense that the first focuses on the development of an infrastructure for managing, sharing and analyzing data as a general service for the entire climate modeling community, while the second is a science application seeking answers to specific questions, such as the accurate modeling of the carbon cycle, with a new generation of simulation code that couples land, sea, geo- and bio-chemistry models within a unified environment.

In a late September 2006 conference call, representatives from VACET and these two climate projects discussed a number of issues ranging from climate science needs for new visualization and analytics techniques to the most promising approaches for deploying and delivering software. The list of stakeholder needs in this area is currently undergoing a process of revision and vetting. In summary, VACET will develop tools that support and improve typical climate studies that involve executing, analyzing and comparing large ensembles of climate simulations, where input and model parameters are sampled over 100's or 1000's of simulation runs resulting in 100s of TB of data.

New climate modeling needs that VACET is planning to address include tools in the following five areas. (1) Comparative analysis tools to visualize coherently multiple models, e.g. with spreadsheet-like presentations of 3D visualizations, to develop better understanding of correlation between variations of input parameters and output simulation. We have identified VisTrails as the best first candidate tool in this area that will automate the process of developing sets of effective visualizations and maintain provenance information associated to each model analyzed. (2) Integrate state of the art 3D visualization libraries in the python scripting environment used by CDAT. (3) Integration of multi-resolution data visualization capabilities with production quality rendering tools to combine the advantages of a interactive visualization environment with the possibility to generate high quality illustrations for publication and presentation. (4) Multi-resolution and streaming technology that combines efficient data processing with data movements. The ESG infrastructure relies heavily on the notion of having data distributed over the network. VACET can provide tools that work effectively within the ESG intelligent management of the data flow (e.g. provide the ability to quickly look at a remote model before committing to its entire transfer, or steam the results for a computation in a progressive to allow preemption of mistakes and save computing resources). (5) Feature identification, tracking and search based initially on visual identification of structures of interest followed by a formal characterization their characteristics using topological, statistical and visualization parameters. Tracking schemes will be based on a salient subset of such parameters. (6) Time-varying 3D vector field visualization specialized for presentation of climate data.

Deployment issues include the desire to roll out new capabilities as part of existing, well-accepted climate data analysis infrastructure and applications (CDAT, Ferret, NCL, etc.).

2.2.5 Fusion

To date, VACET has had interactions with the following fusion projects: (1) Center for Extended Magnetohydrodynamic Modeling (S. Jardin, PPPL), (2) Simulation of Wave Interactions with Magnetohydrodynamics (D. Batchelor, ORNL), (3) Applied Partial Differential Equations Center (P. Colella, LBNL and R. Semtaney, PPPL), (4) Full-scale Reactor Modeling (J. Cary, Tech-X). The scope of these interactions range from introductory and a commitment to engage in collaborative activities to work on defining the project's data understanding and visualization needs and priorities.

In late October, Bethel had a conference call with J. Cary to discuss the scope of the VACET and FACETS project. Bethel will attend a FACETS project kickoff meeting to be held in Boulder, CO in December 2006. One early result is that Tech-X will provide a data reader capable of loading NIMROD and M3D data; VACET will begin integration of that reader into its visualization infrastructure.

VACET team members have been in discussions with Don Batchelor's SWIM project to identify requirement areas for fusion simulation. Those needs include: (1) analysis tools capable of deriving topological information from magnetic field line windings; (2) effective visualization techniques for visualizing 2D and 3D particle flows with large numbers of particles (appx. 10^6); (3) visualization tools capable of displaying variable relationships within high dimensional datasets; (4) computational steering and monitoring on large compute platforms; (5) feature detection at varying spatial and temporal scales.

R. Semtaney, who is part of APDEC, uses an AMR-based solver to model hydrogen pellet injection, which provides the fuel for burning plasmas. Here, AMR helps address challenges posed by large spatial and temporal scale that would not otherwise be possible to solve using a traditional structured grid. His needs include: (1) visualization support for "mapped" AMR grids (i.e., curvilinear adaptive grids); (2) comparative visual analysis between simulation-simulation and simulation-experiment; (3) multigrid AMR mapped grids.

S. Jardin has reported that their existing visualization software infrastructure, which they find quite useful, will not scale to larger datasets. Additionally, their team is not in a position to extend the existing visualization software to support new types of visual exploration or comparative analysis.

The paragraphs above are not a complete representation of stakeholder needs. Instead, they provide an overview of the breadth and depth of their needs. To be successful, VACET will look for areas where stakeholder needs overlap to reach the broadest possible audience.

2.2.6 Applied Partial Differential Equations Center

VACET has a close relationship with the Applied Partial Differential Equations Center. APDEC, which focuses on software infrastructure for PDE solvers on adaptive grids, has science stakeholders in fusion, astrophysics, and combustion, and collaborations with cosmology, hydrology, oceanography, porous media flow and accelerator. By working closely with APDEC to deliver visualization software infrastructure, VACET can achieve impact on a larger number of science stakeholders than would otherwise be possible.

In the first round of SciDAC, APDEC created and delivered its own visualization application (ChomboVis). The high-level objective in SciDAC2 is to replace ChomboVis with new software infrastructure from VACET that is capable of processing large, complex, time-varying AMR datasets created by the APDEC team and that is capable of running on parallel, remote resources.

Several from the VACET team attended the APDEC kickoff meeting in late October 2006 at LBNL. From that meeting, and a couple of follow-up discussions, we have generated a preliminary list of APDEC visualization requirements. That list is currently undergoing a revision and vetting process. Generally speaking, VACET infrastructure needs only a few additional features to meet a baseline set of APDEC needs. We expect this work to be one of VACET's first early successes.

2.2.7 Resources – LBNL/NERSC and ORNL/LCF

As one of VACET's primary objectives is deployment of software infrastructure at DOE's open computing facilities, our team will need access to such resources.

One form of access will come through a formal allocation process. In October 2006, Bethel submitted a formal ERCAP request to NERSC for storage and compute cycles for the VACET team. The first year's request was for 50TB of storage and 20,000 CPU hours. We expect both of these request levels will grow dramatically in out year renewals.

Less formally, VACET team members include staff who are on production visualization teams at both LBNL/NERSC and ORNL/LCF. As staff, they have access to those facilities for the purpose of providing software and consulting services to their customers. Such staff accounts are typically limited in terms of resources – Center staff typically do not consume vast amounts of storage or cycles.

Sean Ahern has arranged for VACET access to the visualization cluster at ORNL/LCF for the purpose of software development and deployment. While this approach will suffice in the short term, it does not grant our team access to the large, leadership class machines. It is likely that VACET will need to submit an INCITE application or Director's Discretionary allocation to access the compute resources at ORNL.

2.2.8 Other Resources

As with any computational research and development project, VACET requires local resources to conduct preliminary investigations prior to rolling out production-quality software at the large centers. Now that the project is officially underway, each institution is in the process of assessing what it needs in terms of local resources to fulfill its obligations on the project. We expect that all institutions will require some amount of additional local resources – compute capacity, storage, graphics hardware – to enable it to conduct its work. The next steps in this area are (1) to identify what resources are needed to do our work and (2) find the funding to pay for those resources.

2.2.9 Begin Initial Deployment

After accounts are in place, we'll be in a position to begin deploying our software infrastructure at NERSC, LCF and on stakeholder platforms. Some amount of interaction with staff at NERSC and LCF may be required in order to install software. Since the VACET teams includes representatives from both NERSC and LCF, we expect no problems in that regard.

2.2.10 First Light Campaigns

Once software is installed and operational, we will conduct "first light" campaigns with those stakeholders who have data and are ready to begin operations. These campaigns will include hands-on tutorials that provide show stakeholders how to use VACET software with their data. We expect that some of these datasets will be rather large and will require use of parallel resources at NERSC and LCF and will involve remote delivery of visualization to the stakeholder.

2.2.11 Project Management

On of our first-year objectives will be to revise this Project Management Plan so it contains a more detailed set of milestones and deliverables based upon the needs of our stakeholders, which include science applications (both SciDAC and non-SciDAC but DOE-sponsored projects like INCITE), other Centers and Institutes. Since we are so early in the project at this stage, detailed milestones and deliverables would not accurately reflect our stakeholders' needs and priorities nor our assessment of what we can realistically accomplish. Specifically, we will be looking for areas of overlap across different stakeholders needs and priorities. We will be in a much better position to provide such information approximately six months into the project, or around 1 April 2007, which is midway through the first year of operations.

2.3 Medium-Term Objectives – Fall 2009

Broadly speaking, we can identify the following list as being mid-project objectives:

- Maturing relationships with stakeholders and collaborators. This means we have documented their needs and priorities, have met some of their objectives and are well underway towards meeting some of the more challenging objectives that require a longer amount of time to complete.

- Effective deployment of VACET software at NERSC and LCF. By mid-project, we expect that use and operations of VACET technology computing facilities – VACET software operations will be "routine."

- Maturing VACET technology. Due to careful software engineering practices and effective intra-team communication, VACET software infrastructure is proving effective at meeting our stakeholder needs and as serving as a platform for deploying new visualization and analytics research.
- Visibility and success. VACET team members will have an extensive record of publications that establishes the team as continuing leaders in the field of visualization basic and applied research.
- Scientific impact. Our stakeholders offer examples of scientific advances made possible by VACET technologies.
- Project expansion. The initial VACET budget, which is 55% of what we requested, is increased to support the complete portfolio of activities needed to achieve PetaScale visualization capabilities for our stakeholders.

2.4 Long-Term Objectives – Fall 2011

Broadly speaking, we can identify the following list as being late-term objectives:

- Productive relationships with stakeholders and collaborators. Late in the project, we expect to have met most, if not all, of our stakeholder and collaborator needs that fall within the scope of the VACET charter and budget. We will have established a plan for going forward into SciDAC3 that builds upon the successes of our stakeholders and of VACET itself.
- Scientific impact. VACET technologies are an indispensable part of our stakeholders scientific investigatory practice, and have led to noteworthy scientific accomplishments.

2.5 Metrics

We have defined the following list of metrics to evaluate our success over the course of our project.

- Are we meeting milestones and deliverables?
- Are we deploying software at DOE's Open Computing Facilities that addresses the data understanding needs of our stakeholders?
- Are we actively engaged with our Science Stakeholders?
- Do we have meaningful interactions with other SciDAC Centers and Institute?
- Number and quality of publications. This category will include journal articles, conference proceedings, and web-based material.
- Do we have an active outreach program?

3 Roles of Key Personnel

3.1 LBNL

3.1.1 Bethel

As coordinating PI for VACET, Bethel's duties include overall management of a five-institution project, direct management and supervision of LBNL staff contributing to the effort, and interfacing with DOE on all matters relating to the project. In addition to project management duties, Bethel will serve as a Stakeholder Project leader, translate Stakeholder needs into project deliverables, prioritize work tasks, design and implement visualization and analytics solutions particularly in the areas of remote, distributed and parallel visualization, query-driven visualization and analysis, and latency-tolerant visualization and rendering.

3.2 ORNL

3.2.1 Ahern

Sean Ahern is the VACET PI for Oak Ridge National Laboratory (ORNL). In addition to directly managing the contributions of the personnel at Oak Ridge to VACET, Ahern will serve as a direct customer liaison to users sited at ORNL, as well as users of ORNL's Leadership Computing Facility (LCF). In this role, Ahern will translate Stakeholder needs into VACET deliverables, prioritize ORNL work tasks, design and implement visualization and analytics solutions particularly where they impact users of the LCF. Specific ORNL customers are likely to be Dr. Don Batchelor in fusion simulation and Dr. John Drake in climate simulation.

3.2.2 Ostrouchov

Dr. George Ostrouchov is a statistician working for the Computer Science and Mathematics division of ORNL. In his role in VACET, Dr. Ostrouchov will serve as a liaison to customers in the area of statistical analysis, feature detection, and coupled statistics/visualization deployment. He will directly liaise with users located at ORNL, specifically Dr. Don Batchelor's fusion project. Dr. Ostrouchov also has continuing collaborations with the Scientific Data Management Center for Enabling Technology, allowing for cross-fertilization between the two centers.

3.2.3 Meredith

Jeremy Meredith is a computer scientist working for the Computer Science and Mathematics division of ORNL. He has extensive experience working with distributed parallel visualization and data analysis codes, especially the VisIt visualization system. He will serve as one of the primary code developers for the VisIt system, integrating new analysis capabilities, and customizing data readers for various VACET customers.

3.3 LLNL

3.3.1 Pascucci

Valerio Pascucci is the coordinator for VACET activities at Lawrence Livermore National Laboratory (LLNL) and will have responsibility for direct management and supervision of LLNL scientists involved in the project. In this capacity, Valerio will translate and prioritize application stakeholder needs into LLNL project deliverables, design and implement data analysis and visualization solutions particularly in the areas of robust topological analysis, feature extraction and tracking, linked views, remote, and distributed streaming techniques.

Valerio will also serve as a Technical Point of Contact for the climate modeling application area and will work with Dean Williams to identify VACET deliverables to be enacted in collaboration with the Earth Science Grid and with John Drake to identify VACET technology that can directly support climate modeling science applications.

3.3.2 Childs

Hank Childs is the VACET Chief Software Architect and is responsible for the implementation of the activities defined by the Executive Committee with respect to software deployment and distribution as well as quality control.

Hank Childs is also the main LLNL developer responsible for introducing new features in VisIt, such as full support AMR data to satisfy the needs of the APDEC center and relative science applications.

3.3.3 Bremer

Peer-Timo Bremer is a Computer Scientist at LLNL. He will be the main person responsible for developing robust topological techniques for data analysis, feature extraction and tracking for structured and unstructured data. Peer-Timo will work with scientists in different applications to develop formal topological characterization of features of interest and corresponding implementation of discrete algorithms for their computation while providing guaranteed error bounds. Peer-Timo will also work on the use of streaming techniques for the analysis of large models from petascale science applications.

3.3.4 Bonnell

Kathleen Bonnell is a Computer Scientist at LLNL who has extensive expertise developing scientific visualization techniques and who has a deep understanding of the VisIt visualization platform. Kathleen will be responsible for integrating a number of new visualization components and data readers within VisIt as well separating particular VisIt portions into stand-alone libraries that can be deployed separately in other environments.

3.4 Utah

3.4.1 Johnson

Professor Chris Johnson is the VACET Co-PI. Along with VACET Co-PI Wes Bethel, he will lead the overall VACET effort. In addition, Johnson will help create new visualization techniques for displaying error and uncertainty, as well as new multi-field visualization techniques.

3.4.2 Hansen

Professor Hansen and a graduate student will be working on multi-field visualization and uncertainty visualization, focusing on GPU methods for volume rendering for multi-fields taking uncertainty into consideration.

3.4.3 Silva

Professor Claudio Silva will pursue the deployment of VisTrails in the climate application for the purpose of comparative analysis. He will be working with Valerio Pascucci at LLNL, who is the direct contact with Dean Williams (climate stakeholder).

3.4.4 Parker

Professor Steven Parker will work on adapting Utah's real-time ray tracer to particle-based datasets (stakeholders in fusion, astrophysics, accelerator modeling). He will also be working with the CPES center to implement a web-based visualization tool to automate routine visualization tasks for the two-dimensional XGC simulation code. Using Web 2.0 technologies, we will enable viewing large numbers of particles and mesh points in a normal web browser.

3.4.5 Sanderson

Dr. Allen Sanderson's role in VACET will be two-fold: (1) a liaison between VACET and fusion and astrophysics scientists addressing their needs as a stakeholder; (2) developing and deploying visualization and analysis tools based primarily on the needs of fusion and astrophysics stakeholders. During the first year, Sanderson will focus primarily on deploying GPU and volume rendering techniques for visualizing and analyzing particle-based datasets.

3.4.6 Tricoche

During the first year, Dr. Xavier Tricoche will focus on multi-field and flow visualization algorithms that target the visual understanding needs of multiple stakeholders. He will contribute software implementations of flow visualization topology-based algorithms in SCIRun and VisIt. He will also work towards the characterization and the representation of salient structures in high-dimensional spaces resulting from the embedding of multiple related quantities in a single analysis space.

3.4.7 Cole

Martin Cole is a staff software engineer who will work on transitioning visualization research algorithms to end user software.

3.5 UC Davis

3.5.1 Joy

Joy's duties involve the interaction with stakeholders to identify fundamental visualization research and development problems, to design and translate these problems into workable solutions, and to develop deliverables that can be returned to our stakeholders to address their data analysis needs. His initial efforts will focus on highly-parallel methods for material interface reconstruction, query-driven visualization, and multi-dimensional visualization. These efforts are targeted at stakeholders in combustion, astrophysics, and applied mathematics.

3.5.2 Hamann

Hamann's duties involve investigating topological solutions and multiresolution methods to support analysis and exploration of petascale data sets. He will focus on out-of-core methods and visual segmentation methods, which allows one to represent massive data sets by a relatively smaller number of elements, and implementing these techniques within the frameworks supported by VACET. Through these efforts, he will collaborate with VACET's LLNL component (Pascucci), to adapt and deploy these methods to our VACET stakeholders.

4 Software and Intellectual Property

4.1 Intellectual Property

Broadly speaking, VACET aims to adapt, develop and deploy new technologies aimed at meeting our science stakeholders' data understanding needs. The primary approach we'll use to achieve that objective is to begin with a solid technology base and then extend and deploy it to meet stakeholder needs. There are two such technology bases – the VisIt and SCIRun – both of which use an Open Source license for distribution.

SCIRun is released under a variant of the MIT license, while VisIt is released under a variant of the BSD license. Both licenses allow for the source code of the respective project to be extended by VACET and made available to the open source community. VACET will own the copyright of contributions made by their developers, although these contributions will still be released under the terms of each project's respesctive Open Source license.

We anticipate there may be new technologies that will be deployed in VisIt, SCIRun or both. These may be libraries, components or some other form of software. In principle and in practice, all VACET institutions are strong Open Source proponents. We anticipate that any new technology emerging from VACET will be released either as part of VisIt or SCIRun, and thus under the terms of their respective Open Source licenses. For technologies not released as part of VisIt or SCIRun, our commitment is to use an Open Source license to minimize the intellectual property barriers to its use by a diverse community of users.

4.2 Software Development, Release and Maintenance

VACET's operational plan to use existing applications as the vehicle to deploy new technology directly addresses several issues related to software development, release and maintenance.

First, the fact that both VisIt and SCIRun are already production-quality applications that are extensible and Open Source, VACET's software development time is much less than would be required were we to develop a completely new application "from scratch." These applications provide much of the infrastructure that is very time consuming to develop, namely graphical user interfaces, data loaders, rendering engines and the internal infrastructure that manages data as it flows from reader to filter to mapper to rendering.

Second, both of these applications are backed by existing development and support infrastructure. Utah and LLNL have development, release engineering and support efforts in place for SCIRun and VisIt, respectively. VACET will provide front-line support to its primary stakeholders over the course of a given project.

Some VACET additions will be appropriate for inclusion as part of the "core" part of VisIt or SCIRun while others won't fit quite so neatly into "the core" of the application. This situation leads to an open questions that will be addressed over the first year of our project. Namely, how will VACET release new technologies that don't fit neatly within the core of one of the primary deployment vehicles? We expect the answer will be that the new technology is a library of some type, and that we will be able to release it using an Open Source license through some established distribution center (e.g., SourceForge or the SciDAC Outreach Center). Use of such a distribution vehicle satisfies requirements to log/count downloads, provide the means for tracking bugs, perform ongoing releases and so forth.

Both these strategies – inclusion as part of "the core" of an established, supported visualization application, or as a standalone "library" but released through an established Open Source distribution vehicle – allow for life beyond VACET for any new VACET technologies.

5 Communication and Outreach

5.1 VACET Internal Communication

The VACET Executive Committee holds bi-weekly conference calls. The purpose of these calls is to discuss project management issues, including but not limited to: (1) stakeholder relations; (2) defining and prioritizing VACET technical objectives; (3) inter-institution technical interactions.

We have established two different email lists for use by the entire VACET team and the VACET Executive Committee. These lists are hosted at LBNL and are quite active.

VACET had a project-wide kickoff meeting on 1 November 2006 in conjunction with IEEE Visualization 2006 in Baltimore, MD.

As of the time of this writing, our most pressing task is to finalize the task lists for our first-round stakeholders. To that end, a battery of frequent phone and face-to-face meetings will occur that focus on refining individual stakeholders' needs.

We anticipate holding quarterly all-hands meetings for the first year to ensure a rapid startup and coordination across the entire team.

5.2 Communication with Other SciDAC Projects

As indicated in previous sections, VACET team members have been very active in reaching out to stakeholder projects. Such outreach includes email, phone calls, and presence at their project meetings. Based upon the early success of such outreach, we expect such frequent and productive communication to continue in the future.

Once the VACET start-up phase is well underway, we will invite representatives from stakeholder projects to VACET all-hands meetings.

We expect to have an active presence at the 2007 SciDAC-wide meeting in Boston. Potential roles there will range from participation in the organizing committee, formal presentations and posters.

We need the ability to have role-based collaborative content authoring on the web. This type of capability, which in scope for the SciDAC Outreach Center, would allow us to have "shared work areas" with our stakeholders and the rest of the SciDAC community. At present, we do not have such a capability and are counting on the SciDAC Outreach Center to provide it.

5.3 External Communication

Our strategy for outreach and communication to "the world" includes a broad array of activities: web, press, technical publications, and technical presentations.

VACET established a website at `www.vacet.org`. The website, while evolving, is intended to provide a comprehensive view of our Center's activities to a wide audience (science stakeholders, DOE program office, general public).

Press releases provide a valuable way to reach out to a potentially large audience that spans academia, government and industry. One example of such a press release is the November 3, 2006 HPCWire interview with VACET team member Bethel. These types of activities help not only VACET, but also SciDAC and DOE's technology portfolio as a whole.

VACET team members have a strong record of technical publications. We expect this trend to continue in the future, thereby maintaining VACET's leadership role in the visualization community.

6 Visualization Center and Institute

As described in the SciDAC RFP [1], Centers for Enabling Technologies provide the software infrastructure that enables new scientific discovery. The role of Institutes is somewhat more broad, but includes education and outreach, focal point for research, and so forth.

The VACET mission is quite clear and responsive to the call: provide visualization and analytics software infrastructure that addresses the data understanding needs of scientific stakeholders where needs include the ability to quickly gain insight into large and complex data, the ability to perform interactive visualization of such large data at DOE's large computing facilities, remote visualization since it is impractical to move large datasets around the Internet, and so forth. Our work plan and stakeholder interactions are engineered to achieve this mission.

The following sections outline probable differences between the Center and Institute based upon the SciDAC call and communication with the program office. These differences are from the Center's point of view.

6.1 Joint Outreach Activities

VACET can provide speakers at the Institute's outreach activities. A good example of such an interaction is the SC06 workshop on Ultra-Scale Visualization held on 13 November 2006 in Tampa FL. Three VACET team members gave technical presentations at that workshop.

6.2 Annual, Multi-Day Workshop

The original VACET proposal included a number of outreach activities that focused on providing current and potential future stakeholders the opportunity for intensive, hands-on training. The idea is to provide the opportunity for science stakeholders to learn about visualization in general, VACET software capabilities in particular, and to have the opportunity to have hands-on-keyboard time with a VACET member to apply VACET technology to their data.

VACET's current budget does not allow us to conduct such annual workshops. We estimated direct costs at about $20K per year to cover costs associated with the workshop, travel costs for stakeholders to encourage their participation and travel costs for a few VACET staff. Costs for effort, which were included in the original VACET budget, were estimated at about two person-weeks per VACET attendee.

6.3 Software

VACET enters SciDAC2 with two established, reasonably mature software collections for delivering visualization and analytics infrastructure to stakeholders. VACET plans to use these platforms not only for delivering production-quality software infrastructure, but also as the basis (a platform) to support future visualization and analytics research and development VACET will need to conduct to meet its ongoing objectives.

Naturally, VACET will be looking for opportunities to incorporate new software into these infrastructures as the opportunities arise. We will achieve this objective through team-wide consistent software engineering practices.

[1] http://www.science.doe.gov/grants/FAPN06-04.html

6.4 Shared Research Staff

Since one of the VACET co-Investigators and the Institute PI are at the same Institution, there is opportunity to share research staff and graduate students between the two projects.

7 Risks

Included below is a preliminary itemization of project risks. Each contains an estimation of potential project impact and remediations.

7.1 Principal Investigator

Scenario: One of the PIs leave the project. Likelihood: very low. Impact: high. Remediation: Due to the high degree of communication within the VACET center and the fact that all on the team understand the VACET mission, the appropriate remediation is to have a backup PI from a lab and one from a University. Bethel and Johnson fulfill the roles of lead and backup PI, respectively.

7.2 Chief Software Architect

Scenario: The CSA leaves the project. Likelihood: very low. Impact: high. Remediation: we have designated Marty Cole (Utah) as the backup CSA.

7.3 Current Budget Levels

Scenario: VACET's budget remains fixed at its present level for the entire five years. Likelihood: medium. Impact: Medium-High. When we engineered our proposal, it consisted of several broad but interlocking/interdependent technology areas that are absolutely required to achieve the objective of "PetaScale visualization." If any of them are eliminated, there is significant risk to VACET achieving its objectives. We felt it is in the best interest of our science stakeholders that we retain critical mass in some areas in order to maximize likelihood of impact. The alternative is to uniformly reduce effort across all technology areas. Technically speaking, that approach is viewed as being an even higher-risk – risk is spread to a greater number of project focus areas, thereby increasing overall project risk. Remediation tactics: (1) augment the VACET effort with SAP funding in the out years to pay for one-on-one stakeholder interactions, thereby freeing up some VACET "core funding" to be brought to bear on areas eliminated by the budget reduction; (2) lobby DOE to reinstate our funding based upon initial early successes with science stakeholders.

7.4 SciDAC Outreach Center

Scenario: The Outreach Center's scope does not include something like a GForge server that supports software downloads, bug tracking, email lists, collaborative content management, etc. Likelihood: medium. Impact: medium – the adverse impact to VACET is that VACET staff time would be required to set up the infrastructure to do software downloads, role-based collaborative document/content editing, and so forth. This effort will reduce the amount of time we have to accomplish our primary mission. Remediation: for software downloads, we could probably use SourceForge, but this approach is not viewed with favor. For collaborative document management and distribution, we may be able to use Google's 'Docs and Spreadsheets" (see `http://docs.google.com`), although this isn't an ideal solution. Recommendation: VACET recommends that the Outreach Center be immediately commissioned and activated according to the scope of its proposal.

8 Appendix A – Future PMP Outline

The following subsection contain additional Project Management Plan components.

8.1 Mission

The Visualization and Analytics Center for Enabling Technologies (VACET) focuses on leveraging scientific visualization and analytics software technology as an enabling technology for increasing scientific productivity and insight. Advances in computational technology have resulted in an "information big bang," which in turn has created a significant data understanding challenge. This challenge is widely acknowledged to be one of the primary bottlenecks in contemporary science. Using an organizational model such as VACET, we are well positioned to respond to the needs of a diverse set of scientific stakeholders in a coordinated fashion using a range of visualization, mathematics, statistics, computer and computational science and data management technologies.

The vision of VACET is to adapt, extend, create when necessary, and deploy visual data analysis solutions that are responsive to the needs of DOE's computational and experimental scientists. Our center is engineered to be directly responsive to those needs and to deliver solutions for use in DOE's large open computing facilities. The research and development directly target data understanding problems provided by our scientific application stakeholders. VACET draws from a diverse set of visualization technology ranging from production quality applications and application frameworks to state-of-the-art algorithms for visualization, analysis, analytics, data manipulation, and data management. Our goal is to respond to the urgent needs of the scientific community by providing significant, production-quality technology to aid in data understanding.

8.2 Goals/Objectives

Our main goal is to deploy a variety of data analysis and visualization tools for our science stakeholders. They have diverse data understanding needs, use a variety of computing resources, and are geographically distributed. Additionally, we want to leverage solutions developed and deployed for one stakeholder to many other projects. We address these challenges by using a flexible approach to software development and project management that draws from the diverse strengths of our team. We have collected input from scientists and engineers, both in person and via an extensive on-line survey [2], from a broad set of fields including climate modeling, fusion, combustion, chemistry, astrophysics and others who are committed to working with us and to use the software tools that we will deploy.

We group the needs expressed by the users (see Figure 1) in two main categories: (1) visualization techniques, ranging from classical rendering techniques to the most advanced data streaming and remote data access algorithms for managing extremely large datasets, and (2) analytics techniques, including data exploration, feature extraction, tracking and comparison that aid the scientist in the actual information discovery process.

8.3 Benefits

Summary: provide the software infrastructure needed to perform petascale visual data analysis in support of DOE's science programs on DOE's open computing facilities.

[2] www.sci.utah.edu/scidac survey.html

Techniques

Applications	Technical Point of Contact	Visualization							Analytics						
		Integration of Basic Tools for Visualization and Analysis	Publication Quality Images, Illustrations, and Movies	Project-Wide Visualization Tools	Flow Techniques	Scalable Tools	Remote Data Access and Streaming Techniques	Collaborative Tools	Query Driven Visualization	Statistical Displays	Feature Detection (topological analysis, semantic range queries)	Temporal specific issues (feature tracking, events, ...)	Multiple Fields	Comparative Visualization and Analytics	Uncertainty Visualization
Fusion	Utah ORNL LBNL	●●	●		●	●		●●	●●●	●●●	●●	●●●	●●●	●●●	●
Combustion	LBNL LLNL ORNL	●	●●●		●●●	●●●	●●●	●●●	●●●	●●	●●●	●●●	●●●	●●	●
Accelerator	LBNL	●	●	●	●	●	●●		●●●	●●	●●●	●●●	●●	●●	●
Astrophysics	ORNL LBNL	●	●●	●	●	●●	●●		●●●	●●	●●	●●	●●●	●●	●
Turbulence	LLNL	●	●●●		●	●●●	●●			●	●●●	●●●	●	●	
Climate	LLNL ORNL	●			●	●●	●●		●●●	●●	●●	●●	●●	●●	●●
Environmental Management	LBNL	●	●●	●●		●●●	●●●	●●							

Figure 1: Relationship between application stakeholders and visualization and analytics techniques.

8.4 Problem

To be completed – draw material from the original proposal.

8.5 Approach

To be completed – draw material from the original proposal.

8.6 Team Organization

As described in the VACET proposal, our Center is organized into functional groups to achieve several distinct objectives: (1) facilitate the flow of information between the Center's leadership, personnel, and science stakeholders; (2) to provide the organizational structure needed to ensure oversight and coordinated operations of the Center's collection of activities; (3) to ensure we meet our work deliverables; (4) to gracefully accommodate future growth and respond to changing priorities. The Center's original functional groups are: the Center PIs, the Executive Committee, the External Advisory Board, Research and Development, Chief Software Engineer, Software Development and Support, and Stakeholder Projects. Due to a 45% reduction in budget, we have eliminated the formal External Advisory Board.

Figure 2 depicts the working interaction between these functional groups. The co-Directors (Bethel and Johnson) are responsible for the operation of the entire Center and are the point of contact between the Center and DOE. The Executive Committee (EC), which makes decisions about research, development, stakeholder applications selection and prioritization, and software deployment, is comprised of the participating lab PIs (Bethel, Johnson, Pascucci, Ahern, and Joy).

Given the high priority on delivering useful software to our scientific stakeholders, each lab site will have a primary software engineer whose duty it is to assure software developed and deployed from that site meets the Center software engineering criteria. The set of site engineers is known collectively as the Software Engineering Group (SEG). In addition to developing, testing, documenting and maintaining the Center software, the Software Engineering Group will integrate

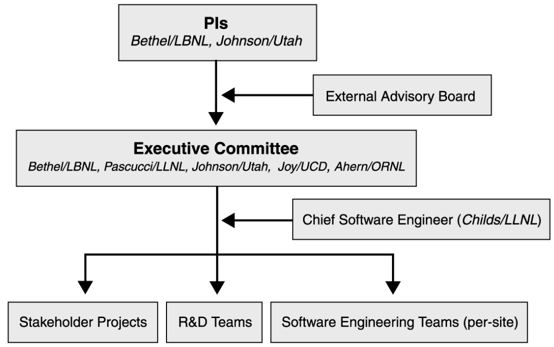

Figure 2: Functional interaction of groups within VACET.

results from all research and development groups into the Center software, relying on feedback from the project leads in the Stakeholder Projects Group, the Executive Committee and the Chief Software Engineer (Childs).

The Chief Software Engineer (CSWE) serves many important functions within the center. One is to facilitate the coordinated design, implementation and integration of the Center's technologies into software solutions that meet stakeholder needs. He will provide guidance to the software development teams so that individual software tools and libraries will be readily usable throughout the Center's collection of stakeholder projects. He will coordinate with the EC to prioritize software development targets, and serve as a technical software advisor to the Center as a whole. He will interact with the R&D team to help foster early designs that fit well within the Center's technology implementations. He will direct the development, testing, deployment, and support of the Center software toolsets.

The Stakeholder Projects Group (SPG) is the primary interface to our science stakeholders. In this group, individuals from the Center will interact directly with science stakeholders to obtain and prioritize science needs, coordinate with the Center's EC and Chief Software Engineer to translate those needs into a work plan, to oversee and manage the work so that software is delivered to the science stakeholder.

Finally, the PIs and EC will work with an External Advisory Board (EAB) to obtain advice/feedback and to update them with status information. The EAB will convene once per year for a one-day meeting to review the Center's progress and strategic plans. The Center will bear the travel costs for the EAB members.

8.7 Roles and Responsibilities

The roles and responsibilities of key personnel were discussed previously in Section 3.

8.8 Interrelationships with Other Projects – Centers, Institutes, Applications

These issues have been discussed with previously in Section 2 (relationships with Science Applications and the APDEC Center), and Section 6 (relationship between the visualization Center and Institute).

8.9 Change Control

TBD.

8.10 Risks

An initial assessment of project risk factors appears previously in Section 7.

8.11 Task Descriptions, Deliverables, Milestones and Schedule

During this "startup" phase of the project, we are focusing on the following preliminary activities: (1) identifying stakeholders, (2) gathering and analyzing stakeholder data understanding needs, (3) vetting such analysis with the stakeholders, (4) translating needs into formal tasks that will result in deliverables, (5) identifying points-of-contact between VACET and collaborators (science applications, other centers and institutes). Once this set of tasks is complete, we will be in a position to generate a more formal list of task descriptions, deliverables, and milestones. Our objective is to provide the first draft of such materials by 1 April 2007.

8.12 Performance Measures

An initial list of project metric appears previously in Section 2.5

8.13 Communications

Communication issues appear previously in Section 5.